DOCUMENT RESUME

ED 426 095		TM 029 325
AUTHOR	Roberts, Debra K.	
TITLE	Partial and Part Correlation Coefficients: Formula and Score Regression Perspectives.	
PUB DATE	1999-01-00	
NOTE	19p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (San Antonio, TX, January 21-23, 1999).	
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)	
EDRS PRICE	MF01/PC01 Plus Postage.	
DESCRIPTORS	*Correlation; *Regression (Statistics); Research Methodology; *Scores	
IDENTIFIERS	Statistical Package for the Social Sciences	

ABSTRACT

Partial and part correlations are discussed as a means of statistical control. Partial and part correlation coefficients measure relationships between two variables while controlling for the influences of one or more other variables. They are statistical methods for determining whether a true correlation exists between a dependent and an independent variable while controlling for one or more other variables. This paper discusses the use and limitations of partial correlations, and presents heuristic data illustrating that computation formulas and regression analyses with latent scores yield equivalent results. Three appendixes contain an example of Statistical Package for the Social Sciences (SPSS) partial correlation computation, SPSS commands for regression analyses, and the SPSS regression analysis printout. (Contains 3 tables, 1 figure, and 10 references.) (SLD)

Partial and Part Correlation Coefficients: Formula and Score Regression Perspectives

Debra K. Roberts

Texas A&M University

Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, January, 1999.

Abstract

This paper discusses partial and part correlations as means of statistical control. Partial and part correlation coefficients measure relationships between two variables while controlling for the influences of one or more other variables. The paper discusses the use and limitations of partial correlations, and presents heuristic data illustrating that computation formulae and regression analyses with latent scores both yield equivalent results.

Partial and Part Correlation Coefficients: Formula and Score Regression Perspectives

Experimental validity is an important concern for researchers (Huck & Cormier, 1996). By implementing methods of experimental and statistical control, better confidence can be placed in one's results. These controls exist in order to reduce extraneous variability in research studies. Some common forms of experimental control include: random assignment, matching of subjects, use of experimental and control groups, and testing alternative hypotheses. While random assignment is considered one of the better variance controls, in most psychological research it is not feasible (Pedhazur, 1982). Statistical control is used as a mathematical means of comparing subjects when they can not be equivalent in fact (McBurney, 1994). Statistical methods are particularly needed when there is more than one independent variable. In these cases, the independent variables are not only related to the dependent variable, but are also related to each other. Researchers need to be able to determine the effects of some variables while controlling others. Partial and part correlations are statistical methods for determining whether a true correlation exists between a dependent and independent variable while controlling for one or more other variables (Waliczek, 1996).

Basic Ideas and Definitions

Partial and part correlation coefficients can be considered Pearson correlation coefficients and are related to multiple regression. In both partial and part correlation, two variables are correlated with the influence of one or more variables partialed out of (i.e., literally removed from) both or only one variable (Hinkle, 1998). For example, a

researcher analyzing the relationship between three variables such as height, weight, and age may want to know if a relationship exists when the age variable is removed. A partial correlation would correlate height and weight while removing the influence of age. A part correlation would correlate height and weight while the age variable is *only removed from one* of the variables.

In order to use partial correlation effectively, the data must be interval or ratio, comparably distributed, and a linear relationship must exist between variables (Korn, 1984; Waliczek, 1996). A partial correlation may be expressed as $r_{xy.z}$; this is interpreted as the correlation between x and y with z removed. Often, numbers are used rather than letters, such as $r_{12.3.}$ A partial correlation with one variable controlled is known as a first-order correlation. More than one variable may be controlled at a time and the coefficients from such analyses are considered second-order, third-order, and so on based on how many independent variables are removed (Pedhazur, 1982). For example, the notation, $r_{12.34}$ implies that this is a second-order correlation with variables 3 and 4 being controlled. Often, a Pearson correlation is called a zero-order correlation due to no variables being controlled since there is only one independent variable.

Examples and Cautions

Partial correlations can be extremely valuable in detecting spurious correlations. For example, this occurs when two variables, such as x and y are correlated due to both being affected by z. However, once z is removed, the correlation between x and y disappears. According to Pedhazur (1982), one of the most common spurious correlations occurs when age is one of the variables. Researchers are not always able to test people of

the same age, therefore, partial correlation assists in determining whether age is contributing to a misleading conclusion. As Pedhazur (1982) noted,

> Partial correlation is not an all-purpose method of control. Its valid application is predicated on a sound theoretical model. Controlling variables without regard to the theoretical considerations about the patterns of relations among them may amount to a distortion of reality and result in misleading or meaningless results. (p. 110)

Figure 1 provides an example of two patterns of causation that are acceptable in performing a partial correlation along with two patterns that are not acceptable. When variables exist with complex patterns of causation, partial correlation should not be used (Pedhazur, 1982; Waliczek, 1996). As stated by Franzblau (1958), partial correlation should not be considered as a substitute for careful research technique and planning.

The formula to compute a first-order partial correlation is:

$$r_{xy.z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1 - r^2_{xz}}\sqrt{1 - r^2_{yz}}}$$

Note that the partial correlation coefficient can be calculated directly from the three correlation coefficients: r_{xy}, r_{xz}, and r_{yz}. The same is true for a first-order part correlation. The formula for a first-order part correlation is:

$$r_{x(y.z)} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1 - r_{yz}^2}}$$

Tables 1 and 2 show the formula calculations for a partial correlation using a

computer spreadsheet program. Table 1 illustrates an extreme case in which the partialed variable *z* causes the correlation coefficient between variables *x* and *y* to increase. Table 2 illustrates another extreme case where the effect is the opposite.

In cases where more than one independent variable is partialed out, the same formulas for partial and part correlations are utilized. For example, if a second-order partial correlation is hand-calculated, one must first calculate three first-order partials to plug into the formula. The formula would appear as:

$$r_{12.34} = \frac{r_{12.3} - r_{14.3}r_{24.3}}{\sqrt{1-r^2_{14.3}}\sqrt{1-r^2_{24.3}}}$$

Another simple approach to computing partial correlation coefficients is by using a computer statistical software program. The Statistical Package for Social Sciences (SPSS) is able to quickly calculate partial correlations. Table 3 contains a small data set and the results from SPSS analysis of these data are illustrated in Appendix A.

As stated earlier, partial and part correlations are related to regression.

> In simplest terms, the partial correlation $r_{12.3}$ is merely the correlation between the residual from predicting X_1 from X_3 and the residual from predicting X_2 from X_3. (Hays, 1994, p. 675)

In regression, the correlation between an independent variable (predictor) and residuals of another variable is always zero. In other words, error scores are always uncorrelated with yhat scores. In a first-order partial correlation with *X*, *Y*, and *Z* variables, two regression equations would be obtained. In partialing out *Z*, we would have one equation for predicting *X* from *Z* and one for *Y* from *Z*. The partial correlation is simply the

correlation of the residuals or error scores of X and Y. An SPSS example is presented here to show that computation formulae and regression analyses with yhat and "e" scores both yield equivalent results.

Using the small data set from Table 3, two regression analyses were performed. Appendix B contains the SPSS commands used in this example. Data for b weights and y-intercepts were obtained and entered into prediction equations. This resulted in a final computer printout, Appendix C, that contains the correlation coefficient between the X "e" scores and the Y "e" scores. By comparing the score regression analyses to the formula computation in Appendix A, we find that the values are identical.

Summary

Partial and part correlations can be usefully applied in many research situations. In fact, statistically controlling for an extraneous variable may even result in a part or partial correlation that is closer to +1 or to –1 (i.e., further from zero) than the original r, if "suppressor effects" are involved (see Henard, 1998).

However, statistical control mechanisms such as these must be used thoughtfully (Franzblau, 1958). A particular problem that can occur, particularly if the control involves quite a few orders (e.g., second-order, third-order, fourth-order) is that it may no longer be clear what is being correlated, after removing so much variance from the two measured variables of primary interest (Thompson, 1995).

References

Franzblau, A.N. (1958). A primer of statistics for non-statisticians. New York: Harcourt Brace and Company.

Hays, W.L. (1994). Statistics(5th ed.). Austin, TX: Harcourt Brace College Publishers.

Henard, D.H. (1998, January). Suppressor variable effects: Toward understanding an elusive data dynamic. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston. (ERIC Document Reproduction Service No. ED 416 215)

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). Applied statistics for the behavioral sciences (4th ed.). Boston, MA: Houghton Mifflin Company.

Huck, S.W., & Cormier, W.G. (1996). Reading statistics and research (2nd ed.). New York: Harper Collins.

Korn, E.L. (1984). The ranges of limiting values of some partial correlations under conditional independence. The American Statistician, 38, 61-62.

McBurney, D.H. (1994). Research methods (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Pedhazur, E.J. (1982). Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York: Holt, Rinehart and Winston.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. Educational and Psychological Measurement, 55, 525-534.

Waliczek, T.M. (1996, January). A primer on partial correlation coefficients. Paper presented at the meeting of the Southwest Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 395 882)

Table 1: Partial Correlation Computation

r_{xy}	r_{xz}	r_{yz}	r_{xz}^2	r_{yz}^2
0.6	0.8	0.1	0.64	0.01

$$\text{Partial1} = \frac{r_{xy} - r_{xz}r_{yz}}{\text{sqr}\ (1-r_{xz}^2)(1-r_{yz}^2)}$$

$$= \frac{0.6-0.8(0.1)}{\text{sqr}(0.36)(0.99)}$$

$$= \frac{0.6-0.8}{\text{sqr}(0.3564)}$$

$$= \frac{0.52}{0.596992}$$

$$= 0.871033$$

Table 2: Partial Correlation Computation

r_{xy}	r_{xz}	r_{yz}	r_{xz}^2	r_{yz}^2
0.75	0.8	0.9	0.64	0.81

$$\text{Partial2} = \frac{r_{xy} - r_{xz}r_{yz}}{\text{sqr}\ (1-r_{xz}^2)(1-r_{yz}^2)}$$

$$= \frac{0.75-0.8(0.9)}{\text{sqr}(0.36)(0.19)}$$

$$= \frac{0.75-0.72}{\text{sqr}(0.0684)}$$

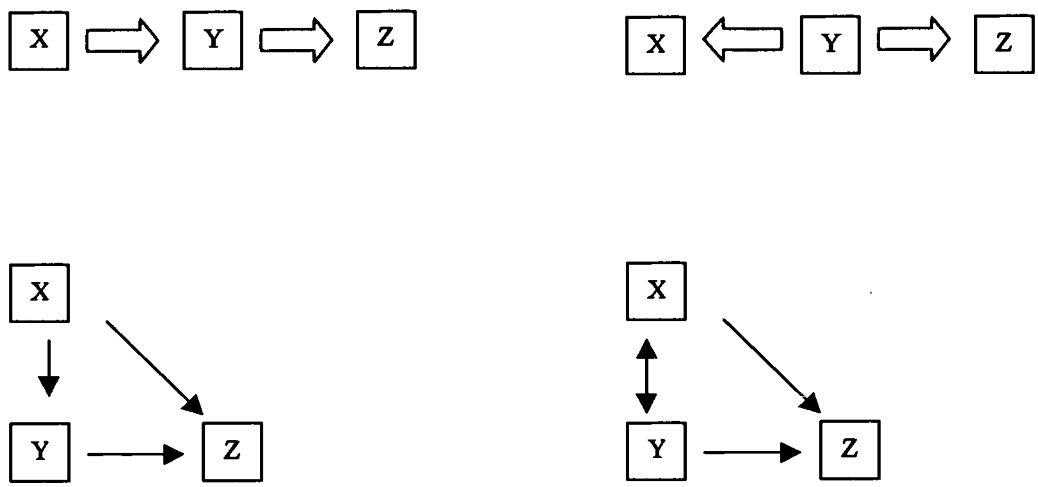
$$= \frac{0.03}{0.261534}$$

$$= 0.114708$$

Table 3: Example Data Set

ID	X	Y	Z
1	15	48	10
2	20	68	34
3	25	36	39
4	30	42	46
5	35	14	17
6	40	60	22
7	45	47	56
8	50	72	60
9	55	80	65
10	60	91	59

Figure 1: Models of Causation



Appendix A

Example SPSS Partial Correlation Computation

```
PARTIAL CORR
    /VARIABLES= var00001 var00002 BY var00003
  /SIGNIFICANCE=TWOTAIL
  /MISSING=LISTWISE .

- -P A R T I A L   C O R R E L A T I O N   C O E F F I C I E N T S - -

Controlling for.. VAR Z

               VAR X       VAR Y

VAR X          1.0000       .2256
              (    0)     (    7)
               P= .        P= .559

VAR Y           .2256      1.0000
              (    7)     (    0)
               P= .559     P= .

(Coefficient / (D.F.) / 2-tailed Significance)
```

Appendix B

SPSS Commands for Regression Analyses

```
TITLE 'Partial correlation'.
COMMENT*************************************.
COMMENT SPSS example.
COMMENT SERA.
COMMENT*************************************.
SET BLANKS=SYSMIS UNDEFINED=WARN PRINTBACK=LISTING.
DATA LIST
     FILE='A:\partial.txt' FIXED RECORDS=1
     /ID 1-2 V1 4-5 V2 7-8 V3 10-11.
EXECUTE.
CORRELATIONS
  /VARIABLES=v1 v2 v3
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE .
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT v1
  /METHOD=ENTER v3 .
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT v2
  /METHOD=ENTER v3 .

compute yhat1=13.986+ (.576*V3) .
compute e1=V1-yhat1 .
compute yhat2=27.034+ (.705*V3) .
compute e2=V2-yhat2 .
LIST VARIABLES=YHAT1 E1/CASES=99 .
LIST VARIABLES=YHAT2 E2/CASES=99 .
correlations variables=V1 V2 V3 yhat1 yhat2 e1 e2/STATISTICS=ALL .
PARTIAL CORR VARIABLES=V1 WITH V2 BY V3 (1) .
```

APPENDIX C

SPSS Regression Analysis Printout

```
09 Nov 98   Partial correlation                                    Page 1

-> DATA LIST
->      FILE='A:\partial.txt' FIXED RECORDS=1
->      /ID 1-2 V1 4-5 V2 7-8 V3 10-11.
-> EXECUTE.
->
-> CORRELATIONS
->   /VARIABLES=v1 v2 v3
->   /PRINT=TWOTAIL NOSIG
->   /MISSING=PAIRWISE .

09 Nov 98   Partial correlation                                    Page 2

                     - -  Correlation Coefficients  - -

             V1          V2          V3

V1           1.0000      .5704       .7466*
V2            .5704     1.0000       .6037
V3            .7466*     .6037      1.0000

* - Signif. LE .05     ** - Signif. LE .01     (2-tailed)          " . " is
printed if a coefficient cannot be computed

09 Nov 98   Partial correlation                                    Page 3
-> REGRESSION
->   /MISSING LISTWISE
->   /STATISTICS COEFF OUTS R ANOVA
->   /CRITERIA=PIN(.05) POUT(.10)
->   /NOORIGIN
->   /DEPENDENT v1
->   /METHOD=ENTER v3 .

09 Nov 98   Partial correlation                                    Page 4

                          * * * *  M U L T I P L E   R E G R E S S I O N
 * * * *

Listwise Deletion of Missing Data

Equation Number 1    Dependent Variable..   V1

Block Number  1.  Method:  Enter      V3

Variable(s) Entered on Step Number  1..    V3

Multiple R            .74663       Analysis of Variance
R Square              .55746                          DF     Sum of Squares     Mean Square
Adjusted R Square     .50214       Regression          1         1149.76456      1149.76456
Standard Error      10.68138       Residual            8          912.73544       114.09193

                                             F =       10.07753       Signif F =   .0131
------------------ Variables in the Equation ------------------

Variable                  B         SE B        Beta          T  Sig T
```

```
V3               .576323      .181547     .746634      3.175  .0131
(Constant)     13.986018     8.140920                  1.718  .1241

End Block Number    1   All requested variables entered.

09 Nov 98   Partial correlation                                  Page 5
-> REGRESSION
->   /MISSING LISTWISE
->   /STATISTICS COEFF OUTS R ANOVA
->   /CRITERIA=PIN(.05) POUT(.10)
->   /NOORIGIN
->   /DEPENDENT v2
->   /METHOD=ENTER v3 .

09 Nov 98   Partial correlation                                  Page 6

                              * * * *   M U L T I P L E   R E G R E S S I O N
 * * * *

Listwise Deletion of Missing Data

Equation Number 1    Dependent Variable..   V2

Block Number  1.  Method:  Enter      V3

Variable(s) Entered on Step Number  1..    V3

Multiple R           .60369        Analysis of Variance
R Square             .36444                         DF     Sum of Squares      Mean Square
Adjusted R Square    .28500        Regression        1         1720.74427       1720.74427
Standard Error     19.36768        Residual          8         3000.85573        375.10697

                                           F =       4.58734       Signif F =  .0646
------------------ Variables in the Equation ------------------

Variable                B          SE B       Beta          T  Sig T

V3               .705050      .329184     .603689      2.142  .0646
(Constant)     27.033973    14.761264                  1.831  .1044

End Block Number    1   All requested variables entered.

09 Nov 98   Partial correlation                                  Page 7
-> compute yhat1=13.986+ (.576*V3) .
-> compute e1=V1-yhat1 .
-> compute yhat2=27.034+ (.705*V3) .
-> compute e2=V2-yhat2 .
-> LIST VARIABLES=YHAT1 E1/CASES=99 .

09 Nov 98   Partial correlation                                  Page 8

   YHAT1        E1

   19.75     -4.75
   33.57    -13.57
   36.45    -11.45
   40.48    -10.48
   23.78     11.22
   26.66     13.34
   46.24     -1.24
   48.55      1.45
```

51.43	3.57
47.97	12.03

Number of cases read: 10 Number of cases listed: 10

09 Nov 98 Partial correlation Page 9
-> LIST VARIABLES=YHAT2 E2/CASES=99 .

09 Nov 98 Partial correlation Page 10

YHAT2	E2
34.08	13.92
51.00	17.00
54.53	-18.53
59.46	-17.46
39.02	-25.02
42.54	17.46
66.51	-19.51
69.33	2.67
72.86	7.14
68.63	22.37

Number of cases read: 10 Number of cases listed: 10

09 Nov 98 Partial correlation Page 11
-> correlations variables=V1 V2 V3 yhat1 yhat2 e1 e2/STATISTICS=ALL .

09 Nov 98 Partial correlation Page 12

Variable	Cases	Mean	Std Dev
V1	10	37.5000	15.1383
V2	10	55.8000	22.9046
V3	10	40.8000	19.6118
YHAT1	10	37.4868	11.2964
YHAT2	10	55.7980	13.8263
E1	10	.0132	10.0705
E2	10	.0020	18.2600

09 Nov 98 Partial correlation Page 13

Variables		Cases	Cross-Prod Dev	Variance-Covar
V1	V2	10	1780.0000	197.7778
V1	V3	10	1995.0000	221.6667
V1	YHAT1	10	1149.1200	127.6800
V1	YHAT2	10	1406.4750	156.2750
V1	E1	10	913.3800	101.4867
V1	E2	10	373.5250	41.5028
V2	V3	10	2440.6000	271.1778
V2	YHAT1	10	1405.7856	156.1984
V2	YHAT2	10	1720.6230	191.1803
V2	E1	10	374.2144	41.5794
V2	E2	10	3000.9770	333.4419
V3	YHAT1	10	1993.8816	221.5424
V3	YHAT2	10	2440.4280	271.1587
V3	E1	10	1.1184	.1243
V3	E2	10	.1720	.0191
YHAT1	YHAT2	10	1405.6865	156.1874
YHAT1	E1	10	.6442	.0716

YHAT1	E2	10	.0991	.0110
YHAT2	E1	10	.7885	.0876
YHAT2	E2	10	.1213	.0135
E1	E2	10	373.4259	41.4918

09 Nov 98 Partial correlation Page 14

- - Correlation Coefficients - -

	V1	V2	V3	YHAT1	YHAT2	E1	E2
V1	1.0000 (10) P= .	.5704 (10) P= .085	.7466 (10) P= .013	.7466 (10) P= .013	.7466 (10) P= .013	.6657 (10) P= .036	.1501 (10) P= .679
V2	.5704 (10) P= .085	1.0000 (10) P= .	.6037 (10) P= .065	.6037 (10) P= .065	.6037 (10) P= .065	.1803 (10) P= .618	.7973 (10) P= .006
V3	.7466 (10) P= .013	.6037 (10) P= .065	1.0000 (10) P= .	1.0000 (10) P= .000	1.0000 (10) P= .000	.0006 (10) P= .999	.0001 (10) P=1.000
YHAT1	.7466 (10) P= .013	.6037 (10) P= .065	1.0000 (10) P= .000	1.0000 (10) P= .	1.0000 (10) P= .000	.0006 (10) P= .999	.0001 (10) P=1.000
YHAT2	.7466 (10) P= .013	.6037 (10) P= .065	1.0000 (10) P= .000	1.0000 (10) P= .000	1.0000 (10) P= .	.0006 (10) P= .999	.0001 (10) P=1.000
E1	.6657 (10) P= .036	.1803 (10) P= .618	.0006 (10) P= .999	.0006 (10) P= .999	.0006 (10) P= .999	1.0000 (10) P= .	**.2256** (10) P= .531
E2	.1501 (10) P= .679	.7973 (10) P= .006	.0001 (10) P=1.000	.0001 (10) P=1.000	.0001 (10) P=1.000	**.2256** (10) P= .531	1.0000 (10) P= .

(Coefficient / (Cases) / 2-tailed Significance) " . " is printed if a coefficient cannot be computed

09 Nov 98 Partial correlation Page 15
-> PARTIAL CORR VARIABLES=V1 WITH V2 BY V3 (1) .

09 Nov 98 Partial correlation Page 16
- - - - - - - - - - - - - - P A R T I A L C O R R E L A T I O N C O E F F I C I E N T S - - - - - - - - - - - - - -

Controlling for.. V3

	V2
V1	**.2256** (7) P= .559

(Coefficient / (D.F.) / 2-tailed Significance) " . " is printed if a coefficient cannot be computed

09 Nov 98 Partial correlation Page 17

TM029325

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Partial and part correlation coefficients: Formula and score regression perspectives

Author(s): Debra K. Roberts

Corporate Source:

Publication Date: 1/99

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
DEBRA K. ROBERTS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here
Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: X Debra K Roberts	Position: RES ASSOCIATE
Printed Name: DEBRA K. ROBERTS	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1831
	Date: 11/9/98